## DETERMINING SPATIAL CORRELATIONS BETWEEN VOTING BEHAVIOR AND SECLECTED DEMOGRAPHIC VARIABLES IN A CHANGING ELECTORATE

Justin Klos Department of Geography, MGIS Program Pennsylvania State University 3600 March Drive, Camp Hill, PA 17011

ABSTRACT: While the United States has existed as a stable two party democracy since its inception, geographically and demographically oriented blocs of voters have often shifted between those parties. Recently in Pennsylvania, political professionals have noted such a voting behavior change in the suburban ring counties around the Commonwealth's two largest cities: Pittsburgh and Philadelphia. The historically conservative counties around Philadelphia are increasingly becoming more politically liberal, while the opposite is occurring in the suburban counties of Pittsburgh as historically liberal counties become more conservative in their voting behavior. This study attempts to statistically and geographically document some of the possible demographic reasons behind this shift in voting behavior by determining geostatistical correlation between demographic variables and voting behavior, as well as attempting to geographically localize this behavior within two study counties: Westmoreland County and Montgomery County. The changes in selected demographic variables between the years 1990-2000 were mapped alongside the change in Democratic voter percentages between the 1988 and 2004 presidential elections. A geostatistical function known as Moran's I analysis was used to determine statistical and spatial correlation between the selected demographic and voting behavior variable pairs, for each municipality in the two county study area. The study discovered several areas where significant voting behavior changes were spatially autocorrelated with changes in demographic factors. The resulting "cluster maps" give the political professional an idea of how and where political behavior has changed.

**Keywords:** Voting behavior, Pennsylvania political demographics, Geostatistical analysis, Spatial autocorrelation, Political geography, Spatial analysis

#### **INTRODUCTION**

Throughout American political history voting behavior has been characterized by long periods of stability punctuated by rapid swings of geographically, demographically or issues-oriented voting blocs from one party to another. While the American two party system has remained relatively stable since the middle of the nineteenth century, the beliefs of American voters have not; shifts of voters between the two parties have been relatively For example, due to a realignment common. centered on civil rights issues between 1956 and 1968 African American vote percentages for the Democratic presidential candidate increased by 34 percentage points, from 60% in 1956 to 94% in 1968 (Campbell & Watson, 2003).

Analysts of Pennsylvania voting behavior believe another significant swing of voting blocs is currently occurring. Analysis of election results over the past twenty years has shown this specifically occurring in the suburban ring counties around the two largest cities in the Commonwealth: Philadelphia and Pittsburgh. Between 1988 and 2004 election returns have showed a steady increase in Democratic presidential vote totals in the Philadelphia suburbs and a similar decrease in Democratic presidential vote totals in the Pittsburgh suburbs. Many explanations have been put forth on why and how this voting behavior is changing, including ideological shifts by voters, in-migration of new voters (in the Philadelphia suburbs) and the rise of a more affluent post-industrial middle class (in the Pittsburgh suburbs) (Madonna, 2006).

The change in voting behavior in these two regions of Pennsylvania have been well documented (Frey & Teixera, 2008; Brownstein, 2008; Madonna, 2006), however potential demographic reasons for the change in behavior or the geographic and spatial characteristics of the behavior on a local basis (i.e. by municipality) have not been thoroughly examined.

The primary research question of this study is to determine whether the changes in voting behavior in the two study counties (Westmoreland outside of Pittsburgh, and Montgomery outside Philadelphia, see Figure 1) are spatially autocorrelated with a corresponding change in demographic variables between 1988 and 2004. In the time scope of the study Westmoreland and Montgomery Counties have exhibited the greatest change in voting behavior of the two suburban rings between 1988 and 2004. The Moran's I Analysis will identify spatially autocorrelated municipalities within the counties. The resulting "cluster maps" of spatially autocorrelated demographic and voting behavior change will allow political campaign professionals to understand how these county and region wide changes in voting behavior and demographics exist spatially within the two target counties.

It is the hypothesis of this research that there will be significant spatially correlated clustering of municipalities within the counties that have high rates of change in the demographic variables and the voting variables. The rest of the article is structured as follows: First a review of the extant literature relevant to the change in voting behavior in the two target counties, Second a quick background on the two study counties, then a presentation of the research methodology and techniques of geostatistical analysis. Finally, the paper will conclude with a discussion of the findings, and a comment on the implications of the study and the benefits of further research on the question.



Figure 1. Study counties: Westmoreland County and Montgomery County.

### LITERATURE REVIEW

The scope of this study, an analysis of the voting behaviors and demographic traits of two particular counties in Pennsylvania, limits the scope of particularly relevant literature. What has been well documented both academically and in the political mass media is that a shift in voting behavior is taking place in the study counties. As the 2008 election neared it was noted that Westmoreland County had voted for the Democratic candidate for

President every time since 1932 (with the exception of 1972) until 1996, but since 2000 it has voted Republican, while conversely Montgomery County voted Republican in every presidential election from 1920 to 1988 (with the exception of the Johnson landslide of 1964); since 1992 it has voted Democratic (Brownstein, 2008). Additionally it has been recognized by political scientists that ideological shifts among the populations of the suburban ring counties and variables such as population in-migration in the Philadelphia suburban counties from the more politically liberal city of Philadelphia and increasing affluence in the Pittsburgh suburban counties may be affecting voter behavior (Madonna, 2006; Frey & Teixera, 2008). It has been firmly established that a change in voting behavior has taken place within the two study counties. What has not been done and what this research attempts to explore is if this change is spatially autocorrelated with change in other demographic variables and where within the two county study area the change is occurring.

The method used here, bivariate spatial autocorrelation considers the strength of the spatial relationship between two variables (in this case a demographic variable and a voting variable) by identifying non-random spatial correlations between those two variables (Sridharan et al., 2007). The when mapped, show clusters results, of municipalities within the two study counties that have had similar changes in both a demographic variable and voting behavior between 1988 and 2004. This clustering effect is most succinctly explained by Tobler's First Law of Geography: "...near things are more related than distant things" (Tobler, 1970; Haining, 2003). The function known as Moran's I analysis is used to determine bivariate spatial autocorrelation, the result of which is a scatter plot showing both overall global autocorrelation and the autocorrelation for each individual geographic unit (Anselin, 2003; O'Sullivan & Unwin, 2002). Conducting Moran's I analysis and determining spatial autocorrelation will allow the hypothesis that voting behavior change and demographic change are related to be tested, and it will also allow us to spatially represent where that change is taking place.

# A BRIEF BACKGROUND ON THE STUDY COUNTIES

Westmoreland County lies directly to the east of Allegheny County and the City of Pittsburgh. Throughout much of the twentieth century it was a center of industry: mining coal, producing coke, and rolling steel in mills along the Monongahela River along the western border of the county. Between 1980 and 2000 40% of manufacturing jobs and 50% of coal mining jobs were lost and the county's population was reduced by 2.74% (United States Census Bureau, 2000). The county has been described as "...old coal and steel country where ethnic lodges, union halls, and loyalty to the Democratic Party once went hand and hand." (Kohut et al., 2005). Westmoreland County began voting for the Democratic presidential candidates in 1932 at the height of the Great Depression and continued to do so with the exception of the Nixon landslide of 1972 until 2000. (Figure 2).



Figure 2. Presidential election vote % 1988-2004, Westmoreland County.

Across the state in the western suburbs of Philadelphia, Montgomery County is a much different case. The heart of what is known as "The Main Line", Montgomery County is composed of prosperous older suburbs directly abutting Philadelphia (such as Lower Merion, Ardmore and Bala Cynwyd), a few industrial towns in the Schuylkill valley (Norristown, Conshohocken), and rapidly developing exurbs in the western part of the county. The population of the county increased 49% between 1960 and 2000, and Montgomery is now Pennsylvania's third most populous county (United States Census Bureau, 2000). Between the years 1989 and 1999 per capita income in the county increased by 40.5%. As the county grew and prospered its political complexion began to change as well. Through much of the twentieth century Montgomery was the keystone in the solid bulwark of Republican counties surrounding Philadelphia. This changed in 1992, and it has voted increasingly Democratic in every election since (Kohut et al., 2005) (Figure 3).



Figure 3. Presidential election vote % 1988-2004, Montgomery County.

#### **METHODOLOGY**

While the broad outlines of the change in voting behavior in these two counties is apparent, the purpose of this study is to examine these changes within the context of changing voting behavior and demographic variables, and to identify spatial patterns within the counties that would indicate geographically where the changes are taking place. A range of demographic variables were selected from the 1990 and 2000 US censuses. The selected demographic variables for each county are as follows:

- Change in Median Household Income 1990-2000
- Change in Median Home Values 1990-2000
- Change in College Degree Educational Attainment 1990-2000
- Change in number Residents migrating from Outside the County in the previous five years 1990-2000
- Change in Population described as living in an "Urban" area 1990-2000
- Change in Percentage of Population age 65+ 1990-2000
- Change in African-American Population

These demographic "change" variables (all rendered as percentages) were derived from the 1990 and 2000 United States censuses. While the voting behavior figures are from 1988 and 2004, the only way to obtain accurate municipal-level demographic data is to use the two most recent censuses. Therefore, while the change in voting behavior is from 1988 to 2004, the change in demographic data is necessarily from 1990 to 2000.

The first step in the analysis was to investigate how the demographic variables and

voting behavior changed on a municipal level within each county, and then representing this change visually using a GIS program. (Figures 4 and 5). Subsequently, the changes in voting behavior and demographic characteristics could begin to be statistically analyzed. Moran's I analysis used as a tool to determine bivariate spatial autocorrelation is defined as:

$$I = \frac{N}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij} (X_i - \bar{X})(X_j - \bar{X})}{\sum_i (X_i - \bar{X})^2}$$

Where there are N geographic units (in this case the number of municipalities in a study county), the attribute value for each unit i is xi, and wij is the weight (or contiguity) for units i and j. (Anselin, 2003; O'Sullivan and Unwin, 2002). In simpler terms. Moran's I creates a spatial matrix of objects (in this case municipalities within the study counties) and measures the degree of relationship between each the values of each object (in this case the variable pairs) in this matrix to determine the statistical correlation between them. A positive spatial autocorrelation indicates non-randomness and a clustering of similar values, while negative spatial autocorrelation indicates non-randomness and a clustering of dissimilar values. In this case we are interested in positive spatial autocorrelation, or determining what municipalities have a high value of change in one variable and neighbors who have high values of change in another variable. The Moran's I analysis was carried out using the GeoDa computer application created by Luc Anselin of the Spatial

Analysis Laboratory, University of Illinois-Champaign. GeoDa allows users to run a variety of standard statistical analyses on spatial lattice data (data that is distributed into spatially discrete units, such as the municipalities in the two target counties). Before the analysis can be conducted a spatial weights index must be created that determines the contiguity between the discreet units (municipalities) in the spatial lattice. In this study the weights index was created using "rook contiguity," meaning units were considered contiguous if they were immediately contiguous and shared any part of their border with the neighboring municipality. The second step in the analysis involves selecting two variables: one original variable and a second or "spatial lag" variable. The Moran's I analysis will measure the relationship between the original variable in each municipality in relation to the spatial lag variable. This required two analyses, the first with voting behavior change as the original variable and the demographic change as the spatial lag, which identifies positive spatially autocorrelated of areas that have high voting behavior change and have neighbors that have high demographic change (Positive spatial autocorrelation between the two variables). The second analysis used the demographic change as the original variable and placed the voting behavior as the spatial lag variable, this analysis in turn identifies positive spatially autocorrelated areas that have had high demographic variable change AND have neighbors with high voting behavior change.



Figure 4. Example of voting behavior change map.



Figure 5. Example of demographic variable change map.

The results of the analysis are displayed in two Moran's I scatter plots. The municipalities (signified by asterisks) in the upper right hand quadrant of each scatter plot are positively autocorrelated (Figure 6). It is then necessary to determine which municipalities have displayed positive spatial autocorrelation in both analyses, i.e. municipalities that have *high values of voting behavior change and high values of demographic variable change and also have*  neighbors who have similar values of voter behavior change and demographic variable change.

The result of this analysis when portrayed geographically gives us clusters of municipalities that have high change rates in each variable (Figure 7). This bivariate Moran's I analysis is repeated for each value of Democratic voter increase or decrease, and increase of each demographic variable value.



Figure 6. Two scatter plots portraying the autocorrelation between Democratic Vote Change (DVOTECHNG) and Median Housing Values (MEDHVPCTCH) in Montgomery County 1988-2004. The variable on the y axis is the spatial lag variable (W\_DVOTECHNG and W\_MEDHVPCTCH), while that on the x axis is the original variable. The upper right hand quadrant contains municipalities that are positive spatially autocorrelated.



Figure 7. Cluster Map Example. Municipalities displaying positive spatial autocorrelation according to Moran's I, between increases in median home value and increase in Democratic presidential vote percentages.

# DISCUSSION OF STUDY RESULTS

In the two county study area twenty-two clusters of spatially autocorrelated demographic/voting variable sets were identified: eleven in each county. While cluster maps were created for each analysis, the limitations of space prohibit them from all being displayed. The mean values of increase and/or decrease for each variable and the municipalities contained within the cluster or clusters for the two county study area can be found in Tables 1 and 2.

Table 1. Mean Values and Locations of Montgomery County Spatially Autocorrelated Clusters

Demographic	Dem Vote Mean	Demographic Mean	Cluster Municipalities	
Variable	Increase	Increase		
African-American	20.29%	4.00%	Collegeville, Hatfield, L. Frederick Twp, L.Merion Twp, L. Providence	
Pop.			Twp, Narberth, Perkiomen Twp, U. Merion Twp	
Median Home	20.01%	20.79%	Collegeville, L. Frederick Twp, L. Merion Twp, L. Providence Twp,	
Value			Narberth, Perkiomen Twp, Schwenksville, Whitemarsh Twp	
Senior Citizen Pop.	19.85%	2.31%	Ambler, Bryn Athyn, E.Norriton Twp, Hatfield, Horsham Twp, L.	
			Moreland Twp, U. Dublin Twp, U. Moreland Twp, W. Norriton Twp,	
			Whitpain Twp	
Outside County	19.51%	6.70%	Bridgeport, Bryn Athyn, Cheltenham Twp, Hatboro, Hatfield,	
Migration			Jenkintown, L. Merion Twp, L. Providence Twp, Perkiomen Twp,	
			Schwenskville, U. Merion Twp.	
College Deg.	19.49%	10.97%	L. Frederick Twp, L. Merion Twp, L. Providence Twp, Narberth,	
Attainment			Perkiomen Twp, Trappe	
Median HH	19.32%	5.70%	Collegeville, Hatfield, L. Frederick Twp, L. Merion Twp, L. Providence	
Income			Twp, Narberth, Perkiomen Twp, U. Merion Twp.	
Poverty Rate	18.93%	-0.33%	Abington Twp, Collegeville, Jenkintown, L. Frederick Twp, U. Dublin	
			Twp, Whitpain Twp	
Urbanization	18.53%	56.88%	L. Frederick Twp, Perkiomen Twp.	
Outside County	-1.74%	9.31%	Green Lane	
Migration				
Senior Citizen Pop.	-4.85%	-3.13%	Towamencin Twp	
Median HH	-7.95%	3.10%	Green Lane, Towamencin Twp	
Income				

Table 2. Mean Values and Locations of Westmoreland County Spatially Autocorrelated Clusters

Demographic	Dem Vote	Demographic	Cluster Municipalities	
variable	Mean Increase	Mean Increase		
Outside County	-18.67%	4.65%	N. Belle Vernon, E. Vandergrift	
Migration				
Urbanization	-18.48%	86.82%	Allegheny Twp, E. Vandergrift, Hyde Park, Oklahoma, W. Leechburg	
Median HH	-15.61%	19.43%	Jeanette, Loyalhanna Twp, N. Belle Vernon, Penn, Rostraver Twp, Smithton, S.	
Income			Huntingdon Twp	
Median Home	-15.50%	27.98%	Fairfield Twp, N. Huntingdon Twp, Sewickley Twp, Smithton, S. Huntingdon	
Value			Twp	
College Deg.	-15.49%	7.45%	Fairfield Twp, Hyde Park, Jeanette, L. Burrell, Seward, St. Clair Twp, Trafford,	
Attainment			W. Leechburg	
Senior Citizen	-15.09%	2.49%	Allegheny Twp, Bell Twp, Bolivar, Fairfield Twp, Hyde Park, Lowe Burrell,	
Pop.			New Florence, Penn Twp, Salem Twp, Seward, Sewickley Twp, St. Clair Twp,	
			Sutersville, W. Leechburg	
Poverty Rate	-12.86%	-5.46%	Mount Pleasant, New Florence, New Kensington, Seward, St. Clair Twp	
Urbanization	0.16%	100.00%	Youngstown	
Median Home	0.16%	22.49%	Youngstown	
Value				
College Deg. Attain	ment 0	.16%	14.20%	Youngstown
Outside County	0	.16%	11.80%	Youngstown
Migration Mean				
Increase				

Overall, in Montgomery County there were eight spatially autocorrelated clusters of municipalities where increases in Democratic voting percentages were spatially autocorrelated with increases of various demographic variables. Two clusters showed a decrease in Democratic voting percentages and an increase in the demographic variable, and a single cluster showed a decrease Democratic vote percentages and a decrease in the demographic variable. NOTE: A reduction in poverty rate is considered an "increase" in that variable for study purposes.

In Westmoreland County there were seven spatially autocorrelated clusters of municipalities where decreases in Democratic voting percentages were spatially autocorrelated with increases of various demographic variables. Four clusters (all four clusters being the single municipality of Youngstown) showed very slight increases in Democratic voting percentage and increases in the demographic variable.

In Montgomery County the correlated variable pairs that showed the greatest increases were Democratic Vote Increase/Median Home Value Increase, Democratic Vote Increase/College Degree Attainment Increase and Democratic Vote Increase/Urbanization Increase. In Westmoreland County the correlated variable pairs that showed the greatest increases were Democratic Vote Decrease/Urbanization Mean Increase, Democratic Vote Decrease/Median Home Value Increase, and Democratic Vote Decrease/Median Household Income Increase.

What is to be made of these results? The study did prove that there are areas of spatial autocorrelation between many of the variable sets in the two county study area, however the old maxim: "Correlation is not causation" must be remembered. No claim can be made that voting behaviors have been influenced by a change in the selected demographic variables or vice versa. Looking at the study results as a whole, however does begin to reveal some patterns.

Many of the spatially autocorrelated clusters of variable pairs in each county cover the same municipalities. While the study cannot claim that high rates of change in demographic variables lead to high rates of change in voting behavior, the cluster maps do indicate that in areas where multiple demographic variables are substantially changing there is a spatially correlated change in voting behavior. Increase in Urbanization and increase in Median Home Values were responsible for two of the three largest increases in voting behavior in each county in the study area. In Westmoreland County an increase in urbanization and home values is

spatially correlated with a decrease in Democratic voting behavior, while in Montgomery County the opposite is the case. There are opportunities to further investigate how these two variables (urbanization and home value increases) affect voting behavior, in particular why they are correlated with two opposite voting behavior trends in two different counties. Indeed, in seven cases a demographic variable increase was spatially autocorrelated with a Democratic voting percent increase in Montgomery County and a voting percent decrease in Westmoreland County. Further research would be required to begin to determine why an increase in a particular demographic variable would lead to an increase in Democratic voting percentages in one county and a decrease in another.

An additional opportunity for further study would involve an expansion of the study's scope. Nationwide election results show that the voting behavior change occurring in Montgomery and Westmoreland County is not unique; there are several examples of similar trends in other metropolitan areas. Further study of counties located in these areas could determine if the spatial autocorrelation of the variable sets is a phenomenon limited to these areas of Pennsylvania, or a phenomenon in voting behavior occurring on national basis.

In addition to expanding our general knowledge on patterns and changes in voting behavior, the appearance of such geographically distinct clusters of spatially autocorrelated voting behavior changes and demographic changes as were found in this study may be of use to political or campaign professionals. In this particular case these statistically proven correlations could help facilitate more accurate voter targeting, but the methods of this study could be applied to any electoral or demographic variables. Political outreach and messaging could be targeted by "cluster" areas. For example, a cluster of municipalities with high home value increases and high democratic voter increases could be more effectively targeted by Democratic candidates with a low property tax political message. In campaign politics at the state and county level the correlation of these factors has always been known or assumed as conventional wisdom, but with this geostatistical function it is possible to document these correlations in a more concrete and statistically robust manner. While a campaign manager might have anecdotal evidence that the Schuylkill Valley is voting more Democratic and new houses are being constructed or being purchased by people with college degrees, this type of analysis makes it possible to statistically quantify the data. And concrete data in an election campaign environment is a very powerful weapon.

## BIBLIOGRAPHY

Anselin, L. 1988. *Spatial Econometrics: Methods and Models*, Boston: Kluwer Academic Press.

Anselin, L. 2003. *GeoDa 0.9 User's Guide*, Retrieved 12 November, 2006 from: www.sal.uiuc.edu/stuff/stuff-sum/pdf/geoda093.pdf.

Barone, M. 1998. *State Legislative Elections: Voting Patterns and Demographics*, Washington, D.C.: Congressional Quarterly.

Brownstein, R. 2008, August 2. Pennsylvania: Schizophrenic State. *National Journal*, 63-65.

Campbell, C. and Watson, R.P. 2003. *Campaigns and Elections: Issues, Concepts, Cases*. Boulder: Lynne Rienner Publishers.

Democratic Office of Demographic Analysis, Pennsylvania House of Representatives. Voting and Election Data.

Frey, W. H. and Ruy, T. 2008. *The Political Geography of Pennsylvania: Not Another Rust Belt State*, Washington, D.C.: The Brookings Institution Blueprint for American Prosperity series.

Haining, R. 2003. *Spatial Data Analysis: Theory and Practice*, Cambridge, U.K.: Cambridge University Press.

Kohut, A. and Taylor, P. and Staff of the Pew Research Center 2005. *Mapping the Political Landscape 2005*, Washington, D.C.: Pew Research Center for the People & The Press.

Lilley, W. 1997. *The State Atlas of Political and Cultural diversity*, Washington D.C.: Congressional Quarterly.

Madonna, T. 2006. *Suburban Tipping Point*, Franklin & Marshall University Center for Politics & Public Affiairs, Retrieved on August 11, 2007 from <u>http://www.fandm.edu/x13075.xml</u>.

Morris, J.D., Roberts, M.S., and Baker, G.F. 1999. *Emotional Responses of African American Voters to Ad Messages*. In Lynda Lee Kincaid & Dianne G. Bystrom (Ed.) The Electronic Election: Perspectives on the 1996 Campaign Communication. New Jersey: Lawrence Erlbaum Associates. Miller, H.J. 2004. *Tobler's First Law and Spatial Analysis*. Annals of the Association of American Geographers, 94 (2), 284-289.

O'Sullivan, D. and Unwin, D.J. 2002. *Geographic Information Analysis*. New Jersey: John Wiley & Sons Inc.

The Pew Research Center 2007. *Trends in Political Values and Core Attitudes (1987-2007)*, Retrieved at <u>http://people-press.org/reports/</u>.

Sridharan, S., Tunstall, H., Lawder, R. and Mitchell, R. 2007. An exploratory spatial data analysis approach to understanding the relationship between deprivation and mortality in Scotland. *Social Science & Medicine*, 65, 1942-1952.

Treadway, J. 2005. *Elections in Pennsylvania*, University Park: Pennsylvania State University Press.

U.S. Census Bureau. 1990. Socioeconomic/Demographic Data. Accessed through the Penn State Pennsylvania State Data Center at: http://pasdc.hbg.psu.edu/pasdc/PA\_Stats/pa\_stats.ht ml.

U.S. Census Bureau. 2000. Socioeconomic/Demographic Data. Accessed through the Penn State Pennsylvania State Data Center at: http://pasdc.hbg.psu.edu/pasdc/PA\_Stats/pa\_stats.ht ml.

U.S. Census Bureau, 2000. County and Municipal boundary data (TIGER/LINE files).